Journal of Biotechnology 251 (2017) 156–159

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Short Genome Communications

Whole-genome *de novo* sequencing of wood rot fungus *Fomitopsis palustris* (ATCC62978) with both a cellulolytic and ligninolytic enzyme system

Chang-Young Hong, Su-Yeon Lee, Sun-Hwa Ryu, Myungkil Kim*

Division of Wood Chemistry & Microbiology, Department of Forest Products, National Institute of Forest Science, Seoul 02455, Republic of Korea

ARTICLE INFO

Keywords:
Fomitopsis palustris (ATCC 62978)
Whole genome sequence
Endoglucanase
Cellulolytic system
Laccase

ABSTRACT

Fomitopsis palustris is a model brown rot fungus causing destructive wood decay based on the cellulase system. Endoglucanase secreted by *F. palustris* hydrolyzes cellulose in both the crystalline and amorphous form. In this study, whole-genome sequencing was conducted to identify genes related to *F. palustris* cellulose degradation and their functions. We determined the 43-Mb complete draft genome of *F. palustris* (ATCC 62978), comprising 14,592 predicted gene models. Gene annotation provided crucial information about the location and function of protein-encoding genes. Three types of endoglucanases were expressed: endo-1,3-beta-glucanase, endo-1,4-beta-D-glucanase, and endoglucanase. In addition, various ligninolytic enzymes such as laccase, aromatic compound dioxygenase, and aryl alcohol dehydrogenase were expressed in *F. palustris* (ATCC 62978). Colony polymerase chain reaction (PCR) indicated that the *endo*-1,4-beta-D-glucanase gene comprises 732 bp. Optimization of the expression conditions of endoglucanase by real-time PCR revealed that endoglucanase was highly expressed after 7 days in all conditions, which was secreted during the secondary metabolism. Studies for large-scale cellulase production from this fungus and investigation of its ligninolytic system will promote its extensive use in various applications. The genomic information determined herein provides a basis for molecular genetics studies to understand the genome functions of *F. palustris* (ATCC 62978).

Cellulose degradation by wood rot fungi occurs via three types of hydrolytic enzymes: endoglucanases (EGs) (EC 3.2.1.4), cellobiohydrolases (CBHs) (EC 3.2.1.91), and β-glucosidase (EC 3.2.1.21). EGs and CBHs cleave the internal bonds of the amorphous and crystalline form of cellulose, respectively. Finally, β-glucosidase liberates glucose from cellobiose (Wood et al., 1989). These cellulases have many potential applications such as biomass utilization, textile industry, pulp and paper industry, and alcohol fermentation (Bhat, 2000; Himmel et al., 1999). The cellulolytic enzymes of fungi have been extensively studied in model basidiomycetes such as *Trichoderma reesi* and *Phanerochaete chrysosporium* (Ilmen et al., 1997; Saloheimo et al., 1988; Uzcategui et al., 1991). However, there is little information on the cellulolytic enzymes secreted from brown rot fungi, which are well known to cause destructive wood decay based on the cellulase system. Among them, *Fomitopsis palustris* is a model basidiomycetes species with processive endoglucanase, resulting in hydrolysis of cellulose in both the crystalline and amorphous forms. In particular, *F. palustris* is used as a test fungus for evaluating the effects of wood preservatives in Korea and Japan (Yoon and Kim, 2005; Yoon et al., 2008). The processive cellulase system of *F. palustris* challenges the conventional knowledge that brown rot fungi lack CBH. Therefore, the *F. palustris* enzyme system shows great potential for cellulase-based applications.

To date, the information on *F. palustris* registered in the National Center for Biotechnology Information database includes sequences of the endoglucanase, glyoxylate dehydrogenase, and isocitrate lyase (Munir et al., 2001a,b). However, studies on the enzyme system of *F. palustris* have focused mainly on the cellulase system, and the complete genome of *F. palustris* remains unknown. Therefore, to better understand the enzyme system of *F. palustris,* we determined the whole-genome *de novo* sequence of *F. palustris* (ATCC 62978) using Pacific Biosciences (PacBio) sequencing technology. This technology can close gaps in current reference assemblies with long reads and facilitates the reliable discovery of novel genes and annotated genes (Rhoads and Au, 2015).

The genomic DNA of *F. palustris* (ATCC 62978) was extracted from the mycelia by phenol-chloroform following shaking culture for 5 days. Complete genome sequencing was conducted using PacBio Single Molecule Real Time (SMRT) sequencing technology (Pacific Biosciences, Menlo Park, CA, USA). The total library was constructed using the PacBio DNA Template Prep Kit 1.0 (for 3–10 kb). SMRTbell templates were annealed using the PacBio DNA/Polymerase Binding Kit P6. The PacBio DNA Sequencing Kit 4.0 and 8 SMRT cells were used for

* Corresponding author.
E-mail address: mkkim0201@korea.kr (M. Kim).

http://dx.doi.org/10.1016/j.jbiotec.2017.04.009
Received 27 October 2016; Received in revised form 6 April 2017; Accepted 11 April 2017
Available online 19 April 2017
0168-1656/

Table 1
Summary of the genome features of *F. palustris* (ATCC 62978).

Parameter	Value
Total length (bp)	43,856,698
N50	667,718
GC content (%)	55.25
Average gene length (bp)	194,056
Protein coding gene (No.)	14,592

sequencing. SMRT cells (Pacific Biosciences) using C4 chemistry and 240-min movies were captured for each SMRT cell using the PacBio RS II (Pacific Biosciences) sequencing platform by Macrogen (Seoul, Korea).

Long reads (mean subread length: 8932 bp) were produced using the PacBio RS II system (Pacific Biosciences). Whole-genome *de novo* assembly was performed using the FALCON method (FALCON v0.2.1), and 226 contigs were generated with an N50 contig length of 667,718. The complete genome of 43,856,698 bp (43 Mb) was obtained. The GC content of the assembly was 55.25%. To correctly predict gene models from an assembled genome, an accurate transcriptome model is required. For this reason, genome-guided transcriptome assembly was performed. RNA reads were mapped to assembled genome sequence using TopHat(v2.0.13) (Trapnell et al., 2009). Then, the assembled transcriptome sequence was obtained from the resultant BAM file using Trinity(r20140717) (Grabherr et al., 2011). Using the assembled genome sequence and assembled transcriptome sequence data, the Seqping(v0.1.33) (Chan et al., 2017) pipeline was followed to annotate the gene models. This pipeline builds gene prediction models GlimmerHMM(v3.0.4) (Majoros et al., 2004), AUGUSTUS(v3.2.2) (Stanke et al., 2006), and SNAP(released 05/17/2012). Subsequently, these prediction results are combined using the annotation program MAKER (v2.28) (Cantarel et al., 2008) contained in the Seqping pipeline. For additional annotation, the consensus sequences were searched against the GenBank non-redundant (NR) database using blast x(v2.4.0+) (Camacho et al., 2009). Finally, 14,592 gene models were predicted (Table 1).

According to the gene annotation, *F. palustris* (ATCC 62978) contains many genes encoding cellulolytic and ligninolytic enzymes, including endo-1,3-beta-glucanase (1 gene), endo-1,4-beta-D-glucanase (1 gene), and endoglucanase (2 genes) (Table 2). In addition, ligninolytic enzymes were annotated in *F. palustris* (ATCC 62978), including cytochrome P450 oxidoreductase (2 genes), laccase (1 genes), aromatic compound dioxygenase (2 genes), and aryl alcohol dehydrogenase (5 genes) (Table 2).

Studies on *F. palustris* enzymes have mainly focused on endoglucanase (Shimokawa et al., 2008; Yoon and Kim, 2005; Yoon et al., 2008). As mentioned above, *F. palustris* was reported to secrete an endoglucanase that cleaves the bonds of both crystalline and amorphous cellulose. In contrast, the ligninolytic enzymes of *F. palustris* have been little studied. However, a preliminary analysis confirmed laccase expression in *F. palustris* (ATCC 62978) based on polymerase chain reaction (PCR) analysis, indicating that the laccase gene contains approximately 1000 bp. Therefore, this PCR result and the present whole-genome sequencing results supported that *F. palustris* (ATCC 62978) has both cellulolytic and ligninolytic enzymes. The annotated ligninolytic enzymes show good ability to degrade lignin polymers. In particular, laccase catalyzes C_{α}-C_{β}, O-C_{β} cleavages and the C_{α} oxidation of lignin compounds (Higuchi, 1990, Morohoshi et al., 1987). Until now, *Gloeophyllum trabeum* was the only brown rot fungus reported to have a laccase gene-specific sequence (D'Souza et al., 1996). Therefore, our observation provides new information regarding the complex cellulolytic and ligninolytic enzyme systems of *F. palustris* (ATCC 62978), which may contribute to the extensive application of this fungus.

PCR analysis was performed to identify the expression of the representative cellulolytic enzyme endoglucanase. Locus tag "GENE_00003562" was identified as endo-1,4-beta-d-glucanase (EC 3.2.1.4) based on the annotation results. For PCR, total RNA was extracted from collected mycelia using the NucleoSpin RNA Plant Kit (Macherey-Nagel, Düren, Germany), and synthesized to cDNA with the addition of M-MLV reverse transcriptase (GenDEPOT, Barker, TX, USA). The synthesized cDNA was used for cloning in a *Pichia* expression vector (pPICZαA). Endoglucanase inserted in pPICZαA was identified by colony PCR. After restriction enzyme digestion with EcoRI and XbaI,

Table 2
Annotated cellulolytic and ligninolytic enzymes of *F. palustris* (ATCC62978).

	Enzyme		Locus Tag	Accession ID[a]	Identity percentage
Cellulolytic enzymes	endo-1,3-beta-glucanase	(EC 3.2.1.6)	GENE_00000089	EPS93808.1	83
	endo-1,4-beta-D-glucanase	(EC 3.2.1.4)	GENE_00003562	BAF49602.1	100
	endoglucanase	(EC 3.2.1.4)	GENE_00009996	BAI58030.1	87/82
			GENE_00011663		
Ligninolytic enzymes	Cythchrome P450 oxidoreductase	(EC 1.6.2.4)	GENE_00001068	EPT00572.1	84
			GENE_00007621	EPT05179.1	96
	NADH-cytochrome b5 reductase	(EC 1.6)	GENE_00000237	KZT74688.1	89
	Laccase	(EC 1.10.3.2)	GENE_00007416	EPT01171.1	87
	aromatic compound dioxygenase	(EC 1.13.11.1)	GENE_00000878	KZT71927.1	87
		(EC 1.13.11)	GENE_00002294	EPT05935.1	92
	Aryl alcohol dehydrogenase	(EC 1.1.1.90)	GENE_00000467	KZT63481.1	78
			GENE_00005484	KZT64790.1	88
			GENE_00005510		
			GENE_00012690		
			GENE_00012704		
	Glutathione S transferase	(EC 2.5.1.18)	GENE_00001397	EPT02850.1	96
			GENE_00002797	EPS98419.1	77
			GENE_00005537	EPT05808.1	85
			GENE_00005613	EPS97040.1	91
			GENE_00008626	KZT66674.1	89
			GENE_00009466	EPT05279.1	80
			GENE_00010528	EPT00877.1	83
			GENE_00010530	EPT00881.1	79
			GENE_00010531		
			GENE_00012329	KZT72292.1	92
			GENE_00012650		

[a] Accession ID were labeled based on BLASTX(v2.4.0) result (use NCBI nr(r20150727) DB).